A Note on the Eigenvalues of the Google Matrix

Lars Eldén
Department of Mathematics, Linköping University

December 29, 2003

Report LiTH-MAT-R–04-01

Let $P \in \mathbb{R}^{n\times n}$ be a column-stochastic matrix, i.e. a matrix with non-negative elements that satisfies $e^T P = e^T$, where $\mathbb{R}^{1\times n} \ni e^T = (1 \quad 1 \quad \cdots \quad 1)$. Define

$$A = \alpha P + (1-\alpha) v e^T,$$

where $0 < \alpha < 1$, and $v \in \mathbb{R}^n$ is a vector with non-negative entries that satisfies $e^T v = 1$. Obviously, A is column-stochastic, $e^T A = e^T$, with positive elements.

A matrix of this type occurs in the computation of pagerank for the Google web search engine [1, 10]. The pagerank vector is the (right) eigenvector of A corresponding to the largest eigenvalue in magnitude, which is equal to 1. The corresponding eigenvector has all non-negative elements, and it is the only eigenvector with this property. This can be proved using Perron-Frobenius theory, see e.g. [9, Chapter 8]. Due to the huge dimension of the matrix, probably between 3 and 4 billion (December 2003), the only viable method for computing this eigenvector is the power method, and variations of it [4, 5, 6] The rate of convergence of the power method depends on λ_2, the second largest eigenvalue (in magnitude) of A, see e.g. [2, Chapter 7.3]. In [3] it was shown that $\lambda_2 = \alpha$. This result was strenghened in [7, 8] to that given in Theorem 1 below.

The purpose of the present paper is to give a simple alternative proof of the theorem.

Theorem 1 ([7, 8]) *Let P be a column-stochastic matrix with eigenvalues $\{1, \lambda_2, \lambda_3 \ldots, \lambda_n\}$. Then the eigenvalues of $A = \alpha P + (1-\alpha) v e^T$, where $0 < \alpha < 1$ and v is a vector with non-negative elements satisfying $e^T v = 1$, are $\{1, \alpha\lambda_2, \alpha\lambda_3, \ldots, \alpha\lambda_n\}$.*

Proof. Define $\hat{e}$ to be e normalized to Euclidean length 1, and let $U_1 \in \mathbb{R}^{n\times(n-1)}$ be such that $U = (\hat{e} \quad U_1)$ is orthogonal. Then, since $\hat{e}^T P = \hat{e}^T$,

$$\begin{aligned} U^T P U &= \begin{pmatrix} \hat{e}^T P \\ U_1^T P \end{pmatrix} (\hat{e} \quad U_1) = \begin{pmatrix} \hat{e}^T \\ U_1^T P \end{pmatrix} (\hat{e} \quad U_1) \\ &= \begin{pmatrix} \hat{e}^T \hat{e} & \hat{e}^T U_1 \\ U_1^T P \hat{e} & U_1^T P^T U_1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ w & T \end{pmatrix}, \end{aligned} \tag{1}$$

where $w = U_1^T P\hat{e}$, and $T = U_1^T P^T U_1$. Since we have made a similarity transformation, the matrix T has the eigenvalues $\lambda_2, \lambda_3, \ldots, \lambda_n$. We further have

$$U^T v = \begin{pmatrix} 1/\sqrt{n}\, e^T v \\ U_1^T v \end{pmatrix} = \begin{pmatrix} 1/\sqrt{n} \\ U_1^T v \end{pmatrix}.$$

Therefore,

$$\begin{aligned} U^T A U &= U^T(\alpha P + (1-\alpha) v e^T) U = \alpha \begin{pmatrix} 1 & 0 \\ w & T \end{pmatrix} + (1-\alpha) \begin{pmatrix} 1/\sqrt{n} \\ U_1^T v \end{pmatrix} \begin{pmatrix} \sqrt{n} & 0 \end{pmatrix} \\ &= \alpha \begin{pmatrix} 1 & 0 \\ w & T \end{pmatrix} + (1-\alpha) \begin{pmatrix} 1 & 0 \\ \sqrt{n}\, U_1^T v & 0 \end{pmatrix} =: \begin{pmatrix} 1 & 0 \\ w_1 & \alpha T \end{pmatrix}. \end{aligned}$$

The statement now follows immediately. □

The theorem implies that even if P has a multiple eigenvalue equal to 1, which is actually the case for the Google matrix, the second largest eigenvalue in magnitude of A is always equal to α.

References

[1] Sergey Brin and Lawrence Page. The anatomy of a large-scale hypertextual Web search engine. *Computer Networks and ISDN Systems*, 30(1–7):107–117, 1998.

[2] G. H. Golub and C. F. Van Loan. *Matrix Computations. 3rd ed.* Johns Hopkins Press, Baltimore, MD., 1996.

[3] T.H. Haveliwala and S.D. Kamvar. The second eigenvalue of the Google matrix. Technical report, Computer Science Department, Stanford University, 2003.

[4] S.D. Kamvar, T.H. Haveliwala, and G.H. Golub. Adaptive methods for the computation of pagerank. *Linear Algebra and its Applications*, 2003, to appear.

[5] S.D. Kamvar, T.H. Haveliwala, C.D. Manning, and G.H. Golub. Exploiting the block structure of the Web for computing PageRank. Technical report, Computer Science Department, Stanford University, 2003.

[6] S.D. Kamvar, T.H. Haveliwala, C.D. Manning, and G.H. Golub. Extrapolation methods for accelerating PageRank computations. In *Proceedings of the Twelfth International World Wide Web Conference*, 2003.

[7] A.N Langville and C.D. Meyer. Deeper inside PageRank. Technical report, Department of Mathematics, N. Carolina State University, 2003.

[8] A.N Langville and C.D. Meyer. Fiddling with pagerank. Technical report, Department of Mathematics, N. Carolina State University, 2003.

[9] C.D. Meyer. *Matrix Analysis and Applied Linear Algebra.* SIAM, Philadelphia, 2000.

[10] L. Page, S. Brin, R. Motwani, and T. Winograd. The PageRank citation ranking: Bringing order to the Web. Stanford Digital Library Working Papers, 1998.